

行政院國家科學委員會專題研究計畫 期中進度報告

子計畫三:IPv6 交換器之高速 IPv6 資料流分類方法之研製 (1/3)

計畫類別：整合型計畫

計畫編號：NSC90-2219-E-032-001-

執行期間：90 年 08 月 01 日至 91 年 07 月 31 日

執行單位：淡江大學電機工程學系(所)

計畫主持人：許獻聰

共同主持人：陳文賢

計畫參與人員：莊岳儒, 王永達, 張智凱, 林俊宏

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 5 月 30 日

行政院國家科學委員會補助專題研究計畫成果報告

以網路處理器為基礎之 IPv6 高層交換器之研製(I)

子計畫三：IPv6 交換器之高速 IPv6 資料流分類方法之研製

(I)

計畫類別：整合型計畫

計畫編號：NSC90-2219-E-032-001

執行期間：90 年 8 月 1 日至 91 年 7 月 31 日

計畫主持人：許獻聰

共同主持人：陳文賢

執行單位：淡江大學電機工程學系

中華民國 91 年 5 月 17 日

行政院國家科學委員會專題研究計畫成果報告

以網路處理器為基礎之 IPv6 高層交換器之研製(I)

子計畫三：IPv6 交換器之高速 IPv6 資料流分類方法之研製(I)

計畫編號：NSC90-2219-E-032-001

執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：許獻聰	執行機構及單位名稱：淡江大學電機工程學系
共同主持人：陳文賢	執行機構及單位名稱：中興大學資訊科學學系
計畫參與人員：莊岳儒	執行機構及單位名稱：淡江大學電機工程學系
王永達	執行機構及單位名稱：淡江大學電機工程學系
張智凱	執行機構及單位名稱：淡江大學電機工程學系
林俊宏	執行機構及單位名稱：淡江大學電機工程學系

一、中文摘要

隨著網際網路的日益發展，而其網路的位址不足、網路頻寬不夠以及網路安全等問題也就越來越重要了。目前各國均積極研發下一代網際網路通訊協定（IPv6）以解決現今網路上的種種問題。以網路設備方面來說，目前不但要具有高速、高效的交換功能，還需要能夠支援新一代的網際網路通訊技術。因此，研發高效的新一代網路交換平台實是刻不容緩。

新一代網際網路通訊協定逐漸被推廣使用，對於新一代網路通訊協定技術的發展也由網路平台的建置、通訊應用程式的開發而進入網路設備的研製。本整合型計畫擬提出建立一個高效能以及高速的交換平台，並在此平台上開發高層（第四到七層）資料分流（Flow Classification）以及新一代網路安全之方法（VPN、IP SEC），再加上研發新一代網路管理之通訊協定（COPS、LDAP），以提供一個完整的新一代網際網路交換設備的系統與解決方案。

本計畫「IPv6交換器之高速IPv6資料流分類方法之研製」為群體研究計劃「以網路處理器為基礎之IPv6高層交換器之研製」之子計劃三。本計劃負責在一個以網路處理器（Network Processor）為基礎的IPv6網路交換平台上，進行IPv6網路第四層

資料分流到第七層URL分類的研發。其目的在為新一代網際網路中提供IPv6資料流分類、URL的分類查詢以及Cookie查詢的方法。

Abstract

Recently, due to the Internet is more popular all over the world. There are some problems including the lack of IP address, bandwidth and the network security also become more and more important. At present, there are many countries to devote to develop the next generation Internet protocol (IPv6) to solve above problems in the networks. In network equipments, they are not only required with high speed and performance of switching, but also to support the IPv6 techniques. Therefore, it is necessary to develop the next generation switching platform.

The next generation Internet protocol (IPv6) is gradually used all over the world. The development processes of IPv6 technology are from the design of a switching platform, writing the application software of communication to implement the network equipment. In this integrated project, we will plan to construct a high performance and speed of a switching platform. At the same time, we also plan to develop the flow classification of upper layers (from layer 4 to

layer 7) and the next generation network security methods (VPN, IP sec.) underlying IPv6. Moreover, to propose the next generation network management protocol (COPS, LDAP), we hope to provide a complete resolution for the next generation network switching system.

The sub-project 3 “the design and implementation of IPv6 flow classification algorithms for high-speed IPv6 switches.” is one of the integrated projects “the design and implementation of network processor-based IPv6 high-layer switches”. This sub-project executes the development of flow classification of layer 4 and URL classified lookup in layer 7 under IPv6 switching platform based on a network processor. The object attempts to provide methods of IPv6 flow classification, URL classified lookup and cookie lookup for the next generation Internet.

二、計劃緣由與目的

目前網際網路正以令人難以預料的速度在膨脹，根據統計平均每年Internet的規模就會擴大一倍。尤其值得注意的是，Internet下一階段發展的動力將不僅僅只是個人電腦市場，而是由多個市場共同的來推動，如：個人移動計算設備、網上娛樂服務、網絡設備控制、家電等等。同時多媒體數據流的加入，對數據流真實性的鑒別，以及提供安全性等方面的需求都迫切要求新一代IP協定的出現。

有新一代之IP(Internet Protocol)協定(簡稱IPv6)被提出來解決此一問題。在此同時，基於以往對Internet的使用經驗得知原本IP協定存在著許多不足之處，如其對即時服務、擁塞控制及保密措施等的支援。IPv6為了滿足對IP的需求，將Address field的大小增加為128 bits。這樣的增加後若將IP平均分配到地球的表面上，每平方公尺估計至少也有1,500個addresses。

在資料流分類的應用愈來愈受重視的現在，IPv6表頭(如圖一)提供了一欄位為流量標籤(Flow Label)，為20位元，是用來識別封包屬於哪一流，以起始位址和流量標籤來共同識別是屬於那一個流量等級。以IPv4來說，因為沒有流量標籤，

所以在處理一些非即時性的應用程式，如電子郵件、通常在背景處理即可，IPv6表頭的流量標籤可以識別一連串的封包屬性，並且在同一個資料串流上每一封包的標籤都是一致的。

除了流量標籤(Flow Label)是我們要注意的，還有一欄位為流量等級(Traffic Class)。為8位元，共可分成256種交通流量等級，這欄位定義了封包的特別服務型式，就是可以利用此一欄位來判斷現行封包所應有的服務型式，而且這個欄位的數值也可經由Router去作變更，以達到網管的目的。

對於應用於資料流分類上，我們可去定義一些數值去代表現行網路上的某些服務。例如，Web、Ftp、Telnet、SMTP、VoD、VoIP等等，將這些服務作策略分級(例如VoIP和SMTP作比較，VoIP屬於即時性的服務，所以屬於較高優先等級的服務)，這代表我們可利用IPv6提供QoS(Quality of Service)功能。

除了利用IP表頭去實現資料流分類以及QoS的功能，我們更可以利用第七層的URL以及Cookie來進行更細節分類動作。我們都知道IP lookup指的是藉由IP address的lookup來執行各種不同的動作(如switching與firewall等)，但URL/Cookie lookup是屬於Layer7的資料流分類(如圖二)。Layer3 Switch最多只是經由IP lookup table作到QoS、Flow Control等功能。URL/Cookie lookup指的也是藉由HTTP表頭內容的lookup作出相應的封包轉送和頻寬管理策略。但值得注意的是，IP lookup是屬於來判定一些較簡單的封包轉送或QoS的策略，但是URL/Cookie lookup是Layer7的層次處理，他是將資料流檔案類型所屬的執行策略類別分析後、再做適當的策略處理。

三、研究方法與成果

在目前市面上已經出現了許多第七層交換器的產品，他們的設計架構大都是圖三所示為主。而且大都是以Network Processor的硬體平台來實現快速的URL/Cookie lookup的動作。為了達到快速

的封包比對動作，Network Processor的硬體架構通常會以多個RISC CPU平行處理的連接方式構成，並且擁有很高的工作時脈和快速的記憶存取系統，以應付大量的使用者存取、快速的資料分類和封包轉送策略的執行。在第一年中我們主要的工作內容是收集資料，閱讀後分析整理找出如何實現快速的URL / Cookie lookup的方法。

目前我們學習到的第七層交換器的封包轉送策略，大都會考慮兩個方向，一是CAP (Content Aware Policy) 內容導向策略，簡單來說其策略是經由URL/Cookie lookup後，由檔案格式或Cookie值來決定封包轉送的策略。另一個則是除了CAP外更考慮了Server State，以上都是很合理的策略考慮方式，因此在未來實現第七層資料流分類的過程中，在CAP和Server State間取得最佳之封包轉送執行策略，是我們研究探討的方向之一。

一開始我們是先從Proxy Server的組成架構去學習，了解了Proxy Server的URL字元查表的資料結構和執行方式。並且也綜合其他實現快速字元查表的方法，得知使用Hash Table是最適合的字元查表的資料結構，不論是考慮搜尋以及資料的更新上，Hash的資料結構都有其顯著的優點在。

接著我們考慮以策略細分的方式，把URL以及Cookie對應的Hash Table作分類，將其單一個Hash Table變成兩個獨立的Hash Table。這樣不但可以試著讓多個處理器可以平行的去搜查Hash Table，並且也會減少Hash Table在更新關鍵字或是策略內容時產生碰撞的機率。

經由搜尋字元的流程來分析，因為我們將Hash Table作成了數個獨立的分類，假設現在想要在最短的時間範圍內搜索一個大小為1514Bytes的乙太網路封包中，當然首先會將MAC、IP、TCP表頭與Payload先獨立的分開，然後才開始去作lookup Payload 的動作 (HTTP表頭是包含在Payload內)，如此分工才有可能使得速度加快。lookup目的是找出所有符合策略的關鍵字元組。關鍵字通常見的是檔案附檔名，當然也可以是路徑名稱、或者完整的檔案名稱、甚至是Cookie的使用者名稱及密碼

指令等，考慮在不同關鍵字的搜尋之下，由於關鍵字位置的前後不同，相對的搜尋的速度也會有所不同。

四、結論與討論

達到快速的URL和Cookie的lookup是我們第一年的目標，接下來一年中的研究過程中我們將會試著藉由IQ2000的發展平台來實現快速的第七層查表動作，藉由充分控利IQ2000架構中的Hash查表結構，並且訂出最適合的分類關鍵字以種類數量，目的是能夠在實驗平台上執行我們的分類演算法，當然我們仍會不斷的學習新知，相信不久的未來一定可以看見豐碩的成果的。

五、參考文獻

- [1] P. Newman et al., "IP Flow Management Protocol Specification for IPv4," *IETF RFC* 1953, May 1996.
- [2] Valeria Cardellini University of Roma Tor Vergata Roma, Italy 00133, Michele Colajanni University of Modrna Modrna, Italy 41100, Philip S. Yu IBM T.J. Watson Research Center "Redirection algorithm for load sharing in Distribute Web server systems" (IEEE1999).
- [3] D.D. Clark, S. Shenker, and L. Zhang, "Supporting Real-time Application in An Integrated Services Packet Network," *Proc. ACM SIGCOMM*, Comp. Commun. Review 22(4), Sep. 1992, 14-26.
- [4] R. Braden, D. Clark, and S. Shenker, "Integrated Services in The Internet Architecture: An Overview," *IETF RFC* 1633, Jul, 1994.
- [5] A. S. Thyagarajan, S. L. Casner, and S. E. Deering, "Making The Mbone Real," *Proc. INET, Honolulu*, Jun. 1995, 465-473.
- [6] Valeria Cardellini University of Roma Tor Vergata Roma, Italy 00133, Michele Colajanni University of Modrna Modrna, Italy 41100, Philip S. Yu IBM T.J. Watson Research Center "Dynamic load balancing in geographically distributed heterogeneous Web servers" (IEEE1998).
- [7] RFC 1953 I P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. Liaw, T. Lyon, G. Minshall, "Ipsilon Flow Management Protocol Specification for IPv4 Version 1.0," 05/23/1996.

[8] RFC 1987 I P. Newman, W. Edwards, R. Hinden, E. Hoffman, F. Liaw, T. Lyon, G. Minshall, "Ipsilon's General Switch

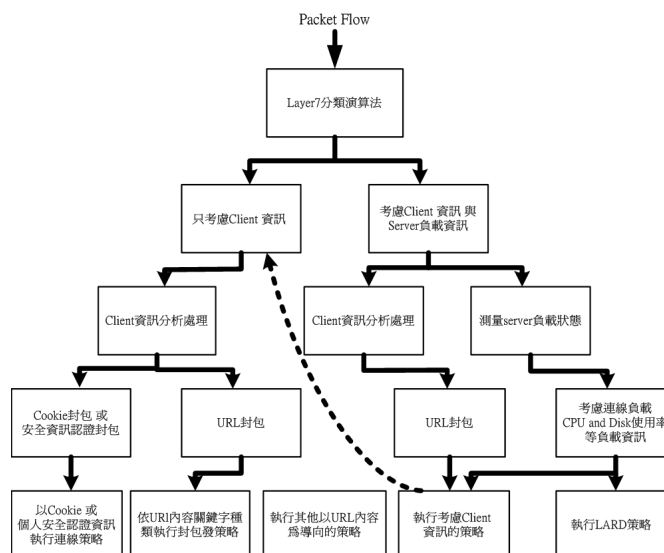
Management Protocol Specification Version 1.1," 08/16/1996.

六、圖表

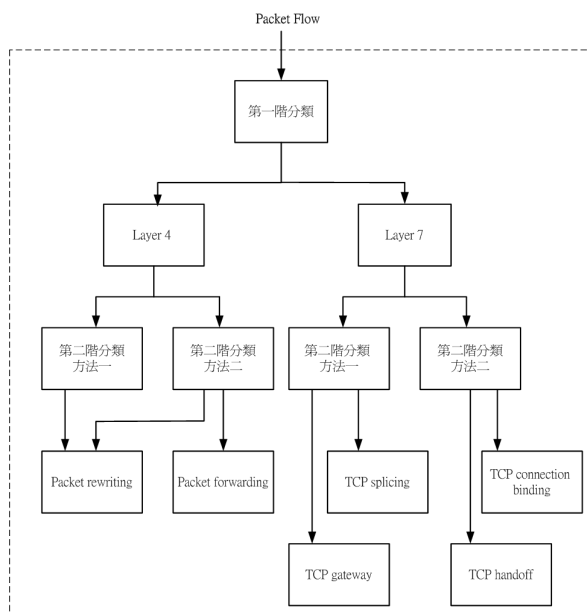
Ipv6 表頭

Ver	Traffic Class	Flow Label	
Payload Length		Next Header	Hop Limit
Source IP Address			
Destination IP Address			
Data Portion of Datagram			
Payload			

圖一、Ipv6 固定欄位格式。



圖二、第七層封包資料分類流程圖。



圖三、一般第七層網交換器架構圖。